AD-785 623

ON THE ARITHMETIC MEANS AND VARIANCES OF PRODUCTS AND RATIOS OF RANDOM VARIABLES

Fred Frishman

Army Research Office
Durham, North Carolina

1971

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AD 785 623

ON THE ARITHMETIC MEANS AND VARIANCES OF PRODUCTS AND RATIOS OF RANDOM VARIABLES

FRED FRISHMAN, DR.
U. S. ARMY RESEARCH OFFICE
DURHAM, NORTH CAROLINA

1. INTRODUCTION

In the sciences and other disciplines, it is quite common to encounter situations where random variables appear in combinations. Two combinations are the product and the ratio of two random variables. Given such combinations, it is required to know the mean and the variance and it is preferable that the mean and variance be determined from the original random variables. Thus, if $U = XY$ and $V = X/Y$ then we want to know the arithmetic means and variances of U and V based on our knowledge of X and Y.

2. ON THE PRODUCT OF TWO RANDOM VARIABLES

The mean and the variance of a product are well known [5]. To obtain the result in a more simple manner, we note that

(2.1) $$\mathrm{Cov}(X,Y) = E(XY) - E(X)\,E(Y)$$

provided that the moments exist. Henceforth, we will assume that the moments exist.

By rearranging (2.1), we obtain

(2.2) $$E(XY) = \mathrm{Cov}(X,Y) + E(X)\,E(Y).$$

Further, the definition of the linear correlation coefficient $\rho_{X,Y}$ is given by

(2.3) $$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{[\mathrm{Var}(X)]^{1/2}[\mathrm{Var}(Y)]^{1/2}}$$

16

By introducing (2.3) into (2.2), it is seen that

$$(2.4)\qquad E(XY) = E(X)E(Y) + \rho_{X,Y}[\mathrm{Var}(X)]^{1/2}[\mathrm{Var}(Y)]^{1/2}\,.$$

To obtain the variance of a product of random variables, we note that

$$(2.5)\qquad \mathrm{Var}(XY) = E(X^2Y^2) - [E(XY)]^2,$$

and that

$$(2.6)\qquad \mathrm{Cov}(X^2,Y^2) = E(X^2Y^2) - E(X^2)E(Y^2).$$

Then from (2.1) and (2.5),

$$\begin{aligned}\mathrm{Cov}(X^2,Y^2) &= \mathrm{Var}(XY) + [\mathrm{Cov}(X,Y) + E(X)E(Y)]^2 - E(X^2)E(Y^2)\\ &= \mathrm{Var}(XY) + [\mathrm{Cov}(X,Y)]^2 + 2\,\mathrm{Cov}(X,Y)\,E(X)E(Y)\\ &\quad - \mathrm{Var}(X)\,\mathrm{Var}(Y) - \mathrm{Var}(X)\,[E(Y)]^2 - \mathrm{Var}(Y)\,[E(X)]^2.\end{aligned}$$

Then, by rearrangement of terms,

$$\begin{aligned}(2.7)\qquad \mathrm{Var}(XY) &= \mathrm{Cov}(X^2,Y^2) - [\mathrm{Cov}(X,Y)]^2 - 2\,\mathrm{Cov}(X,Y)\,E(X)E(Y)\\ &\quad + \mathrm{Var}(X)\,\mathrm{Var}(Y) + \mathrm{Var}(X)\,[E(Y)]^2 + \mathrm{Var}(Y)\,[E(X)]^2.\end{aligned}$$

If X and Y are independent or uncorrelated, (2.2) and (2.7) reduce to

$$(2.8)\qquad E(XY) = E(X)\,E(Y),$$

and

$$(2.9)\qquad \mathrm{Var}(XY) = \mathrm{Var}(X)\,\mathrm{Var}(Y) + \mathrm{Var}(X)\,[E(Y)]^2 + \mathrm{Var}(Y)\,[E(X)]^2.$$

In addition, if the two variables are independent or uncorrelated, then

$$(2.10)\qquad \mathrm{Var}(XY) = \mathrm{Var}(X)\,\mathrm{Var}(Y)$$

only if the means of both random variables are zero.

3. ONE SET OF RESULTS FOR THE MEAN AND VARIANCE OF A RATIO OF RANDOM VARIABLES

We can write, using (2.3)

$$(3.1)\qquad \rho_{Y,1/X} = \frac{E\left(\frac{Y}{X}\right) - E(Y)\,E\left(\frac{1}{X}\right)}{[\mathrm{Var}(Y)]^{1/2}\left[\mathrm{Var}\left(\frac{1}{X}\right)\right]^{1/2}}\,.$$

Then as in (2.4),

$$(3.2)\qquad E\left(\frac{Y}{X}\right) = E(Y)\,E\left(\frac{1}{X}\right) + \rho_{Y,1/X}[\mathrm{Var}(Y)]^{1/2}\left[\mathrm{Var}\left(\frac{1}{X}\right)\right]^{1/2}$$

provided that all the terms exist. In Section 5, it is shown that $E[\frac{1}{X^r}]$ exists provided that the density function has a zero of order at least r at the origin. For example, $E[\frac{1}{X^r}]$ does not exist for the Normal Distribution.

We obtain $\mathrm{Var}\left(\frac{Y}{X}\right)$, by replacing X with $\frac{1}{X}$ in (2.7).

$$(3.3)\quad \mathrm{Var}\left(\frac{Y}{X}\right) = \mathrm{Cov}\left(Y^2, \frac{1}{X^2}\right) - [\mathrm{Cov}(Y, \frac{1}{X})]^2 - 2\mathrm{Cov}\left(Y, \frac{1}{X}\right)E(Y)\,E\left(\frac{1}{X}\right)$$
$$+ \mathrm{Var}(Y)\,\mathrm{Var}\left(\frac{1}{X}\right) + \mathrm{Var}(Y)[E(\frac{1}{X})]^2 + \mathrm{Var}\left(\frac{1}{X}\right)[E(Y)]^2.$$

If Y and $\frac{1}{X}$ are uncorrelated, or if Y and X are independent (see Section 5), then

$$(3.4)\quad E\left(\frac{Y}{X}\right) = E(Y)\,E\left(\frac{1}{X}\right)$$

and

$$(3.5)\quad \mathrm{Var}\left(\frac{Y}{X}\right) = \mathrm{Var}(Y)\,\mathrm{Var}\left(\frac{1}{X}\right) + \mathrm{Var}(Y)\,E\left(\frac{1}{X}\right)^2 + \mathrm{Var}\left(\frac{1}{X}\right)[E(Y)]^2,$$

under the usual assumptions of existence.

If Y and X are uncorrelated Normals, $E\left(\frac{Y}{X}\right)$ and $\mathrm{Var}\left(\frac{Y}{X}\right)$ do not exist. The proof of the lack of existence of the mean is shown in Section 5. Since the mean does not exist, neither does the variance.

In all of the above, we note the requirement for information about the reciprocal of the denominator term.

4. ANOTHER RESULT FOR THE MEAN AND VARIANCE OF A RATIO OF RANDOM VARIABLES.

By using (2.3) we can write

$$(4.1)\quad \rho_{X,Y/X} = \frac{E(Y) - E(X)\,E\left(\frac{Y}{X}\right)}{[\mathrm{Var}(X)]^{1/2}\left[\mathrm{Var}\left(\frac{Y}{X}\right)\right]^{1/2}}$$

Then, by rearrangement of terms,

$$(4.2)\quad E\left(\frac{Y}{X}\right) = \frac{E(Y)}{E(X)} - \frac{\rho_{X,Y/X}[\mathrm{Var}(X)]^{1/2}\left[\mathrm{Var}\left(\frac{Y}{X}\right)\right]^{1/2}}{E(X)}$$
$$= \frac{E(Y)}{E(X)} - \frac{\mathrm{Cov}\left(X, \frac{Y}{X}\right)}{E(X)},$$

provided that the terms exist.

(Also from (2.7), where we replace Y by $\frac{Y}{X}$,

$$(4.3)\quad \mathrm{Var}(Y) = \mathrm{Cov}\left[X^2, \left(\frac{Y}{X}\right)^2\right] - \left[\mathrm{Cov}\left(X, \frac{Y}{X}\right)\right]^2 - 2\,\mathrm{Cov}\left(X, \frac{Y}{X}\right) E(X)\, E\left(\frac{Y}{X}\right) + \mathrm{Var}(X)\,\mathrm{Var}\left(\frac{Y}{X}\right) + \mathrm{Var}(X)\left[E\left(\frac{Y}{X}\right)\right]^2 + \mathrm{Var}\left(\frac{Y}{X}\right)[E(X)]^2.$$

Rearrangement gives

$$(4.4)\quad \mathrm{Var}\left(\frac{Y}{X}\right) = \frac{\mathrm{Var}(Y) - \mathrm{Cov}\left[X^2, \left(\frac{Y}{X}\right)^2\right] + \left[\mathrm{Cov}\left(X, \frac{Y}{X}\right)\right]^2}{\mathrm{Var}(X) + [E(X)]^2} + \frac{2\,\mathrm{Cov}\left(X, \frac{Y}{X}\right) E(X)\, E\left(\frac{Y}{X}\right) - \mathrm{Var}(X)\left[E\left(\frac{Y}{X}\right)\right]^2}{\mathrm{Var}(X) + [E(X)]^2}.$$

Introducing (4.2) into (4.4), we get

$$(4.5)\quad \mathrm{Var}\left(\frac{Y}{X}\right) = \left\{\mathrm{Var}(Y) - \mathrm{Cov}\left(X^2, \frac{Y^2}{X^2}\right) - \left[\mathrm{Cov}\left(X, \frac{Y}{X}\right)\right]^2 + 2E(Y)\,\mathrm{Cov}\left(X, \frac{Y}{X}\right) - \frac{\mathrm{Var}(X)}{[E(X)]^2}\left[E(Y) - \mathrm{Cov}\left(X, \frac{Y}{X}\right)\right]^2\right\} \frac{1}{\mathrm{Var}(X) + [E(X)]^2}$$

If X and $\frac{Y}{X}$ are uncorrelated,

$$(4.6)\quad E\left(\frac{Y}{X}\right) = \frac{E(Y)}{E(X)} \text{ provided that } E(X) \neq 0,$$

and

$$(4.7)\quad \mathrm{Var}\left(\frac{Y}{X}\right) = \frac{\mathrm{Var}(Y)\,[E(X)]^2 - \mathrm{Var}(X)\,[E(Y)]^2}{[E(X)]^2\,\{\mathrm{Var}(X) + [E(X)]^2\}},$$

provided that $E(X) \neq 0$.

Analagous with Section 3, we are highly dependent on information about relationships between the numerator random variable and the ratio of the random variables.

5. DETERMINATION OF THE MEAN OF A RATIO BY USE OF CONDITIONAL PROBABILITIES

It follows from the definition of conditional probability that

$$(5.1)\qquad E\left[\frac{Y}{X}\right] = E_Y\left\{E_{X|Y}\left(\frac{Y}{X}\Big|Y\right)\right\}$$

$$= E_Y\left\{[Y]\ E_{X|Y}\left(\frac{1}{X}\Big|Y\right)\right\} .$$

If we consider, as a special case the Bivariate Normal, we know that the marginal distributions including the conditional are Normal distributions. And for the Normal, it is well known that $E[\frac{1}{X}]$ does not exist so that $E[\frac{Y}{X}]$ does not exist.

More generally, if X and Y are independent, then

$$E[\frac{Y}{X}] = E[Y]\ E[\frac{1}{X}]$$

since if Y and X are independent, then Borel functions of Y and X are independent.

The phenomenon of the occurence of the non-existence of $E[\frac{Y}{X}]$ is portrayed more graphically in the case of independence in a paper by H. A. David [2, pp. 122-123] as follows:

Let $G(x)$ be a cumulative distribution function. Define a central moment of negative order $-r$ by

$$(5.2) \qquad \mu'_{-r}(x) = \int_{-\infty}^{+\infty} x^{-r}\ dG(x)\ dx.$$

This moment exists in the discrete case and also in the continuous case provided that the density function $g(x)$ has a zero of order at least r at the origin (i.e., $\lim_{x\to 0} x^{-r} g(x) = 0$).

Consider the case where the random variables are continuous. Let Y be independent of X with density function $f(y)$; and if $\mu'_{-r}(x)$ and $\mu'_r(y)$ exist,

$$(5.3) \qquad \mu'_{-r}(x)\, \mu'_r(y) = \int_{-\infty}^{+\infty} x^{-r} g(x)\ dx \int_{-\infty}^{+\infty} y^r\ f(y)\ dy.$$

or,

$$(5.4) \qquad E\left[\left(\frac{Y}{X}\right)^r\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\frac{Y}{X}\right)^r g(x)\ f(y)\ dx\ dy = E[Y^r]\ E\left[\frac{1}{X^r}\right].$$

In order that the density function $g(x)$ have a zero of order of least r at the origin, we are saying, in effect, that the random variable is either all positive or all negative, and if 0 is in the range of the random variable, then $g(x)$ must be an infinitesimal at $X=0$.

6. RAO'S PROCEDURE FOR DETERMING $E(\frac{Y}{X})$

Rao [7] determines the Expected Value of the ratio of Normally distributed variates in essentially the following fashion:

Let X and Y be bivariate normal $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$; $-\infty < x,\ y < +\infty$.

Now, let $U = \frac{Y-\mu_y}{\sigma_y}$ and $Z = \frac{X-\mu_x}{\sigma_x}$, so that $Y = \mu_y + U\sigma_y$ and $X = \mu_x + Z\sigma_x$. From the definitions of U and Z, we have, $E(U) = E(Z) = 0$ and $\sigma_u^2 = \sigma_z^2 = 1$.

Then,

$$(6.1)\qquad \frac{Y}{X} = \frac{\mu_y + U\sigma_y}{\mu_x + Z\sigma_x} = \mu_y\left(\frac{\mu_y + U\sigma_y}{\mu_y}\right)\frac{1}{\mu_x}\left(\frac{\mu_x}{\mu_x + Z\sigma_x}\right)$$

$$= \frac{\mu_y}{\mu_x}\left(1 + \frac{U\sigma_y}{\mu_y}\right)\left(1 + \frac{Z\sigma_x}{\mu_x}\right)^{-1}.$$

Rao expanded the last term under the assumption of convergence of the sum. It is well known, however, that $\frac{1}{1+a}$ converges provided that $-1 < a < +1$. This requirement in our case is that $-1 < \frac{Z\sigma_x}{\mu_x} < 1$ or equivalently $|Z\sigma_x| < |\mu_x|$. In terms of the original random variable, the requirement is that $-1 < \frac{X-\mu_x}{\mu_x} < +1$ or that $0 < X < 2\mu_x$ for $\mu_x > 0$, and that $2\mu_x < X < 0$ for $\mu_x < 0$. But since $-\infty < X < +\infty$ for the Normal Distribution, the requirement is not satisfied. The procedure requires that $\mu_x, \mu_y \neq 0$, also.

Nevertheless, the procedure is appropriate for random variables which satisfy the convergence conditions so we continue the development of Rao's procedure, where Z and U are appropriate standardized random variables.

We note that finite limits of the denominator random variable are a prerequisite. As examples, the procedure is valid for appropriately parameterized uniform or binomial random variables. It probably is appropriate for cases where we say that the Normal

distribution satisfies the data, since we probably mean that a truncated Normal satisfies the data, it being rare for an item or test measurement to take on both positive and negative values. By slightly rewriting (6.1) and performing the expansion, we obtain

$$(6.2)\qquad \frac{Y}{X} = \frac{1}{\mu_x}\left(\mu_y + U\sigma_y\right)\left[1 - \frac{Z\sigma_x}{\mu_x} + \frac{Z^2\sigma_x^2}{\mu_x^2} - \frac{Z^3\sigma_x^3}{\mu_x^3} + \cdots\right]$$

$$= \frac{\mu_y}{\mu_x}\left[1 - \frac{Z\sigma_x}{\mu_x} + \frac{Z^2\sigma_x^2}{\mu_x^2} - \frac{Z^3\sigma_x^3}{\mu_x^3} + \cdots\right]$$

$$+ \frac{\sigma_y}{\mu_x}\left[U - \frac{ZU\sigma_x}{\mu_x} + \frac{Z^2U\sigma_x^2}{\mu_x^2} \cdots\right]$$

Taking Expected Values, we obtain

$$(6.3)\qquad E\left[\frac{Y}{X}\right] = \frac{\mu_y}{\mu_x}\left\{1 + \frac{\sigma_x^2 E(Z^2)}{\mu_x^2} - \frac{\sigma_x^3 E(Z^3)}{\mu_x^3} + \cdots + \frac{(-1)^m \sigma_x^m E(Z^m)}{\mu^m} + \cdots\right\} - \frac{\sigma_y}{\mu_x}\left\{\frac{\sigma_x E(ZU)}{\mu_x} - \frac{\sigma_x^2 E(Z^2U)}{\mu_x^2} + \cdots + \frac{(-1)^{n-1}\sigma_x^n E(Z^nU)}{\mu_x^n} + \cdots\right\}$$

$$= \frac{\mu_y}{\mu_x}\left\{1 + \sum_{i=2}^{\infty} (-1)^i \frac{\sigma_x^i E(Z^i)}{\mu_x^i}\right\} + \sigma_y \sum_{j=1}^{\infty} \frac{(-1)^j \sigma_x^j E(Z^jU)}{\mu_x^{j+1}}$$

If, in addition to the prior conditions on the range of X and that $\mu_x \neq 0$, we specify that Z, U are bivariate central symmetrically distributed. Then from [8,pp.23], we have that

$$E(Z^{2i+1}) = 0, \quad i = 0,1,\ldots$$

and

$$E(Z^{i}U) = 0 \text{ for } i \text{ an even integer.}$$

Consequently,

$$(6.4) \quad E\left[\frac{Y}{X}\right] = \frac{\mu_Y}{\mu_X}\left\{1 + \sum_{j=1}^{\infty} \frac{\sigma_X^{2j} E(Z^{2j})}{\mu_X^{2j}}\right\} - \sigma_Y\left\{\sum_{j=1}^{\infty} \frac{\sigma_X^{2j-1} E(Z^{2j-1}U)}{\mu_X^{2j}}\right\}.$$

If, in addition, we add the condition of independence between Z and U or X and Y, then (6.4) reduces to

$$(6.5) \quad E\left[\frac{Y}{X}\right] = \frac{\mu_Y}{\mu_X}\left\{1 + \sum_{i=1}^{\infty} \sigma_X^{2i} \frac{E(Z^{2i})}{\mu_X^{2i}}\right\},$$

which exists under the usual conditions.

We note that (6.3), (6.4) and (6.5) presented above could have been obtained by use of a Taylor Expansion of $\frac{Y}{X}$, the classical Propagation of Errors procedure.

We conclude this Section by noting that Rao's procedure for determining $E\left[\frac{Y}{X}\right]$ is valid only under very special circumstance. These are not satisfied by the Bivariate Normal.

7. THE VARIANCE OF $\frac{Y}{X}$, FOLLOWING FROM RAO'S PROCEDURE

Following the classical definition, we can write

$$(7.1) \quad \mathrm{Var}\left(\frac{Y}{X}\right) = E\left[\frac{Y}{X} - E\left(\frac{Y}{X}\right)\right]^2 .$$

For the terms in the square brackets, we utilize (6.2) and (6.3) and after extensive algebraic simplification, we arrive at

$$(7.2) \quad \mathrm{Var}\left(\frac{Y}{X}\right) = \frac{\mu_Y^2}{\mu_X^2}\left\{\frac{\sigma_X^2}{\mu_X^2} + \sum_{i=2}^{\infty} \frac{\sigma_X^{2i}}{\mu_X^{2i}} \mathrm{Var}(Z^{i}) + \right.$$

$$2\sum_{\substack{i=k\\ i<k}}^{\infty}\ \sum_{k=2}^{\infty}\frac{\sigma_x^{i+k}}{\mu_x^{i+k}}\,(-1)^{i+k}\,\mathrm{Cov}(Z^i,Z^k)\Bigg\}$$

$$+\frac{2\mu_y\sigma_y}{\mu_x^2}\Bigg\{(-1)\frac{\sigma_x}{\mu_x}\,\mathrm{Cov}(Z,U)\ +$$

$$\sum_{i=1}^{\infty}\ \sum_{j=1}^{\infty}(-1)^{i+j}\,\frac{\sigma_x^{i+j}}{\mu_x^{i+j}}\,\mathrm{Cov}\,(Z^i,Z^jU)\Bigg\}+$$

$$\frac{\sigma_y^2}{\mu_x^2}\Bigg\{1+\sum_{j=1}^{\infty}\frac{\sigma_x^{2j}}{\mu_x^{2j}}\,\mathrm{Var}(Z^jU)$$

$$+\ 2\sum_{\substack{j=0\\ j<m}}^{\infty}\ \sum_{m=1}^{\infty}\frac{\sigma_x^{j+m}}{\mu_x^{j+m}}\,(-1)^{j+m}\,\mathrm{Cov}\,(Z^jU,Z^mU)\Bigg\}$$

where, for example, $\mathrm{Var}\,(Z^jU)=E[Z^jU-E(Z^jU)]^2$

$$\mathrm{Cov}\,(Z^jU,Z^mU)=E[Z^jU-E(Z^jU)]\,[Z^mU-E(Z^mU)]$$

$$\mathrm{Cov}\,(Z,U)=E[Z-E(Z)]\ [U-E(U)]\ =E(ZU).$$

If we take only the first term in each of the curly brackets, we obtain an approximation

$$\mathrm{Var}\left(\frac{Y}{X}\right)\approx\frac{\mu_y^2\,\sigma_x^2}{\mu_x^4}-2\,\frac{\rho_{uz}\mu_y\sigma_y\sigma_x}{\mu_x^3}+\frac{\sigma_y^2}{\mu_x^2}\tag{7.3}$$

$$=\frac{\mu_y^2\sigma_x^2+\mu_x^2\sigma_y^2-2\,\rho\mu_x\mu_y\sigma_x\sigma_y}{\mu_x^4}$$

which can be obtained, also, by the usual Propogation of Errors procedures.

If we specify that X and Y (and consequently U and Z) are independent, then (7.2) reduces to

$$(7.4)\qquad Var\left(\frac{Y}{X}\right)=\frac{\mu_y^2}{\mu_x^2}\left\{\frac{\sigma_x^2}{\mu_x^2}+\sum_{i=2}^{\infty}\frac{\sigma_x^{2i}}{\mu_x^{2i}}Var(Z^i)+2\sum_{\substack{i=1\\ i<k}}^{\infty}\sum_{k=2}^{\infty}\frac{\sigma_x^{i+k}}{\mu_x^{i+k}}(-1)^{i+k}Cov(Z^i,Z^k)\right\}+\frac{\sigma_y^2}{\mu_x^2}\left\{1+\sum_{j=1}^{\infty}\frac{\sigma_x^{2j}}{\mu_x^{2j}}Var(Z^jU)\right\},$$

and the approximation reduces to

$$(7.5)\qquad Var\left(\frac{Y}{X}\right)\simeq\frac{\mu_y^2\sigma_x^2+\mu_x^2\sigma_y^2}{\mu_x^4}$$

8. THE DISTRIBUTION OF THE RATIO OF NORMAL VARIATES

C. C. Craig [1, pp. 24-32] presents Fieller's development [3, pp. 424-440] on the distribution of a ratio and Geary's approximation [4, pp. 442-460]. The following is a summary of the results:

Assume (X,Y) is Bivariate Normal or BVN $(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$. Then

$$(8.1)\qquad f(x,y)=\frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_x\sigma_y}e^{-1/2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2-2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)+\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}.$$

Let $v=\frac{y}{x}$ and let $x=x$. Then, $dv\,dx=\frac{1}{|x|}dx\,dy$. We define

$$(8.2)\qquad g(v)=\int_{-\infty}^{+\infty}|x|f(x,xv)\,dx=\int_0^{\infty}x\,f(x,xv)\,dx-\int_{-\infty}^{0}x\,f(x,xv)\,dx$$

$$=\int_{-\infty}^{+\infty}x\,f(x,xv)\,dx-2\int_{-\infty}^{0}x\,f(x,xv)\,dx$$

Letting

$Q(v)=\int_{-\infty}^{+\infty}x\,f(x,xv)\,dx$, and $R(v)=\int_{-\infty}^{0}-2x\,f(x,xv)\,dx$, then

(8.3) $$g(v) = Q(v) + R(v)$$

where $R(v) > 0$ since $-2 \int_{-\infty}^{0} x \ f(xv, x)\, dx > 0$.

First, it can be shown that

(8.4) $$Q(v) = \frac{\mu_x \sigma_y^2 - \rho \sigma_x \sigma_y (\mu_y + v\mu_x) + v\mu_y \sigma_x^2}{\sqrt{2\pi}\,(\sigma_y^2 - 2\rho v \sigma_x \sigma_y + v^2 \sigma_x^2)^{3/2}} \cdot e^{-\frac{1}{2}\left\{\frac{(\mu_y - v\mu_x)^2}{\sigma_y^2 - 2\rho v \sigma_x \sigma_y + v^2 \sigma_x^2}\right\}} .$$

Now, $R(v)$ is of the same form as $Q(v)$ with the exception that the limits of integration are different. Thus

(8.5) $$R(v) = \frac{\sigma e^{-\frac{1}{2\sigma^2}\left(\frac{C}{A}\right)}}{\pi\sqrt{A}} - 2Q(v)\,\Phi\left(\frac{B}{A\sigma}\right) ,$$

where $$\Phi\left(\frac{-B}{A\sigma}\right) = \int_{-\infty}^{\frac{-B}{A\sigma}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx .$$

Now, returning to the distribution of the ratio (8.3), $g(v) = Q(v) + R(v)$, we replace $R(v)$ by the result in (8.5) to obtain

(8.6) $$g(v) = Q(v) - 2\,Q(v)\,\Phi\left(-\frac{B}{A\sigma}\right) + \frac{\sigma e^{-\frac{1}{2\sigma^2}\left[\frac{C}{A}\right]}}{\pi\sqrt{A}}$$

$$= Q(v)\left\{1 - 2\Phi\left(-\frac{B}{A\sigma}\right)\right\} + \frac{\sigma e^{-\frac{1}{2\sigma^2}\left[\frac{C}{A}\right]}}{\pi\sqrt{A}}$$

Thus we have the distribution of the ratio in closed form, where $Q(v)$ is given by (8.4), and

$$
(8.7)\quad \left\{
\begin{aligned}
A &= \sigma_y^2 - 2\rho v\sigma_x\sigma_y + v^2\sigma_x^2 \\
B &= \mu_x\sigma_y^2 + v\mu_y\sigma_x^2 - \rho(\mu_y + v\mu_x)\,\sigma_x\sigma_y \\
C &= \mu_x^2\sigma_y^2 - 2\rho\,\sigma_x\sigma_y\mu_x\mu_y + \mu_y^2\sigma_x^2 \\
\sigma^2 &= \frac{(1-\rho^2)\,\sigma_x^2\sigma_y^2}{A}\ ,
\end{aligned}
\right.
$$

we can rewrite $Q(v)$ by using (8.7) as follows:

$$
(8.8)\qquad Q(v) = \frac{B}{\sqrt{2\pi}\,A^{3/2}}\; e^{-\frac{1}{2\sigma^2}\left(\frac{AC-B^2}{A^2}\right)} .
$$

Now, we return to (8.5), and (8.7). We note that as $\mu_x \to \infty$, then $B \to \infty$ and $C \to \infty$. As a result,

$$
(8.9)\qquad e^{-\frac{1}{2\sigma}\,\frac{C}{A}} \to 0\ ,
$$

and

$$
\Phi\left(-\frac{B}{A\sigma}\right) \to 0\ .
$$

From this, we conclude that

$$
(8.10)\qquad R(v) \to 0 \quad \text{as} \quad \mu_x \to \infty\ .
$$

Therefore, for μ_x sufficiently large, we can get a good approximation of $g(v)$ by $Q(v)$, only.

Thus, for the case where μ_x is large, from (8.4) we have

$$
(8.11)\ Q(v) = \frac{\mu_x\sigma_y^2 - \rho\sigma_x\sigma_y(\mu_y + v\mu_x) + v\mu_y\sigma_x^2}{\sqrt{2\pi}\,(\sigma_y^2 - 2\rho v\sigma_x\sigma_y + v^2\sigma_x)^{3/2}}\; e^{-\frac{1}{2}\left\{\frac{(\mu_y - v\mu_x)^2}{\sigma_y^2 - 2\rho v\sigma_x\sigma_y - v^2\sigma_x^2}\right\}}
$$

We impose the transformation

$$
(8.12)\qquad z = \frac{v\mu_x - \mu_y}{(\sigma_y^2 - 2\rho v\sigma_x\sigma_y + v^2\sigma_x^2)^{1/2}}\ , \quad \text{with}\ dz = \frac{B}{A^{3/2}}\, dv,
$$

and we note that the denominator of $Q(v)$ is positive for all v. Hence

$$(8.13) \qquad Q(z) = Q\left[\frac{v\mu_x - \mu_y}{(\sigma_y^2 - 2\rho v \sigma_x \sigma_y + v^2 \sigma_x^2)^{1/2}}\right].$$

It follows that, as $\mu_x \to \infty$ for σ_x finite, z approaches a $N(0,1)$. This is the Geary result [4].

If in (8.6), $\mu_x = \mu_y = 0$, we obtain

$$(8.14) \qquad g_1(v) = \frac{(1-\rho^2)^{1/2}}{\pi\left(\frac{\sigma_y}{\sigma_x} - 2\rho v + v^2 \frac{\sigma_x}{\sigma_y}\right)},$$

a generalized Cauchy distribution.

The cumulative distribution function for (8.14) is

$$(8.15) \qquad G(v) = \int_{-\infty}^{v} g(v)\; dv = \frac{1}{\pi} \text{ arc tg}\left[\frac{\sigma_x v - \rho\sigma_y}{\sigma_y (1-\rho^2)^{1/2}}\right]$$

as obtained from Gradstein and Ryzhik, [6, p. 82].

The generalized Cauchy obtained by this writer is a simple extension and has utility more from a theoretical viewpoint than an applications viewpoint.

However, the Geary result has great applicability as do the approximation results in Sections 6 and 7 when one is willing to assume that X and Y are Bivariate Normal. As noted in Section 6, in the real world, few if any, random variables are truly normally distributed since few, if any, can range from $-\infty$ to $+\infty$. Few, if any, will take on both positive and negative values! Thus the mathematical abstraction, the Normal Distribution, can be at best a good approximation, and a good approximation only around the mean of the data, since ordinarily one does not <u>see</u> data points in the extreme tails of the Distribution.

9. SUMMARY

In Sections 3 and 4, exact expressions are given for the Mean and the Variance of a Ratio of random variables, regardless of the form of the random variables. One problem is that certain correlation coefficients must be known or assumed to be zero in order to use the formulae. Also, there is the usual requirement that the moments exist. Based on the results of Section 5, we know that these exact results are not appropriate for the Normal Distribution and those other distributions where the random variable takes on negative and positive values and/or the value 0 with probability greater than 0.

In Section 6, a procedure is given for getting the Mean of a Ratio. The requirement is that the denominator random variable be such that either $2\mu_x < X < 0$ when μ_x is negative or that $0 < X < 2\mu_x$ when μ_x is positive. Also, we require that the Mean of the numerator $\mu_y \neq 0$. In Section 7, we obtain the Variance. Since the results in both Sections require summing infinite series, approximations can be obtained by terminating the summation wherever desired.

Finally in Section 8, we sketch out the development for an approximation of the distribution of the ratio if both random variables are Normally Distributed, in the sense that the Denominator variable does not take on both positive and negative values. It is claimed that it is rare for actual random variables to be Normally distributed in the mathematical sense. Rather, the Normal Distribution is a reasonable approximation at best, and under these circumstances, the results presented in Sections 6,7 and 8 are appropriate to be used.

10. BIBLIOGRAPHY

1. Craig, C. C. (1942) "On Frequency Distributions of the Quotient and of the Product of Two Statistical Variables", Amer. Math. Monthly, Vol. 49, No. 1.

2. David, H. A. (1955) "Moments of Negative Order and Ratio-Statistics", Journal of the Royal Stat. Society, Part 1.

3. Fieller, E. C. (1932) "The Distribution of the Index in a Normal Bivariate Population", Biometrika, Vol. 24.

4. Geary, R. C. (1932) "The Frequency Distribution of the Quotient of two Normal Variables", Journal of the Royal Statistical Society, Vol. 93.

5. Goodman, L. (1960) "On the Exact Variance of Products", Journal of the American Stat. Assn., Vol. 55.

6. Gradstein, I. S. and Ryzhik, I. M. (1965) Tables of Integrals, Series and Products. Academic Press.

7. Rao, C. R. (1952) Advanced Statistical Methods in Biometrics Research. John Wiley and Sons, Inc.

8. Frishman, F. (1971) "On Ratios of Random Variables and an Application to a Non-linear Regression Problem". Ph.D. dissertation, The George Washington University.